14 Mar 2015

In the Analyzing Animal Societies book (Whitehead 2008) chapter 9.4 (appendices) the method for calculating social differentiation in socprog is explained. Basically, social differentiation S is defined as the coefficient of variation of the association indices, α_{ij}, for individuals ij where $i \neq j$. The coefficient of variation CV is defined as the ratio of the standard deviation to the mean:

$$CV = \frac{\sigma}{\mu}$$

Our association indices are just estimates calculated by the number of times individuals ij are seen together, x_{ij}, divided by the total number of samples for i and j, d_{ij}. We can think of α_{ij} as the true unobserved association index and x_{ij} a random variable binomially distributed.

$$x_{ij} \sim Binom(d_{ij}, \alpha_{ij})$$

The association indices are also treated as random variables ranging between 0 and 1 drawn from the beta distribution.

$$\alpha_{ij} \sim Beta(u_1, u_2)$$

The u_1 and u_2 are the beta distribution parameters. They can be solved in terms of the mean, μ, and standard deviation, σ from which we can easily calculate the social differentiation coefficient S.

$$u_1 = \mu \cdot \left(\frac{1-\mu}{\mu \cdot (\frac{\sigma}{\mu})^2} - 1 \right)$$
$$u_2 = (1-\mu) \cdot \left(\frac{1-\mu}{\mu \cdot (\frac{\sigma}{\mu})^2} - 1 \right)$$

Socprog uses a maximum likelihood monte carlo approach to choose μ and σ, by maximizing:

$$Likelihood = Binom(x_{ij}; d_{ij}, \alpha_{ij}) \cdot Beta(\alpha_{ij}, u_1, u_2)$$

The likelihood is written slightly differently in Whitehead (2008), but it is equivalent. To make this a posterior distribution instead of a likelihood, I added priors for μ and σ. Both of these priors are truncated normal, centered on the estimated μ_{est} and σ_{est} calculated from x_{ij} and d_{ij} and a large variance of 1000. Both distributions are truncated at 0, since negative values don't make sense in the context of association indices.

$$\mu \sim N(\mu_{est}, 1000);\ 0 < \mu < \infty$$
$$\sigma \sim N(\sigma_{est}, 1000);\ 0 < \sigma < \infty$$

I used JAGS (just another gibbs sampler) and R to estimate μ and σ.